

EXHIBIT E



Duplicate content

Duplicate content generally refers to substantive blocks of content within or across domains that either completely match other content or are appreciably similar. Mostly, this is not deceptive in origin. Examples of non-malicious duplicate content could include:

- Discussion forums that can generate both regular and stripped-down pages targeted at mobile devices
- Store items shown or linked via multiple distinct URLs
- Printer-only versions of web pages

If your site contains multiple pages with largely identical content, there are a number of ways you can indicate your preferred URL to Google. (This is called "canonicalization".) [More information about canonicalization.](#)

However, in some cases, content is deliberately duplicated across domains in an attempt to manipulate search engine rankings or win more traffic. Deceptive practices like this can result in a poor user experience, when a visitor sees substantially the same content repeated within a set of search results.

Google tries hard to index and show pages with distinct information. This filtering means, for instance, that if your site has a "regular" and "printer" version of each article, and neither of these is blocked with a [noindex meta tag](#), we'll choose one of them to list. In the rare cases in which Google perceives that duplicate content may be shown with intent to manipulate our rankings and deceive our users, we'll also make appropriate adjustments in the indexing and ranking of the sites involved. As a result, the ranking of the site may suffer, or the site might be removed entirely from the Google index, in which case it will no longer appear in search results.

There are some steps you can take to proactively address duplicate content issues, and ensure that visitors see the content you want them to.

- **Use 301s:** If you've restructured your site, use [301 redirects](#) ("RedirectPermanent") in your .htaccess file to smartly redirect users, Googlebot, and other spiders. (In Apache, you can do this with an .htaccess file; in IIS, you can do this through the administrative console.)
- **Be consistent:** Try to keep your internal linking consistent. For example, don't link to `http://www.example.com/page/` and `http://www.example.com/page` and `http://www.example.com/page/index.htm`.
- **Use top-level domains:** To help us serve the most appropriate version of a document, use top-level domains whenever possible to handle country-specific content. We're more likely to know that `http://www.example.de` contains Germany-focused content, for instance, than `http://www.example.com/de` or `http://de.example.com`.

- **Syndicate carefully:** If you syndicate your content on other sites, Google will always show the version we think is most appropriate for users in each given search, which may or may not be the version you'd prefer. However, it is helpful to ensure that each site on which your content is syndicated includes a link back to your original article. You can also ask those who use your syndicated material to use the [noindex meta tag](#) to prevent search engines from indexing their version of the content.
- **Use Webmaster Tools to tell us how you prefer your site to be indexed:** You can tell Google your [preferred domain](#) (for example, <http://www.example.com> or <http://example.com>).
- **Minimize boilerplate repetition:** For instance, instead of including lengthy copyright text on the bottom of every page, include a very brief summary and then link to a page with more details. In addition, you can use the [Parameter Handling tool](#) to specify how you would like Google to treat URL parameters.
- **Avoid publishing stubs:** Users don't like seeing "empty" pages, so avoid placeholders where possible. For example, don't publish pages for which you don't yet have real content. If you do create placeholder pages, use the [noindex meta tag](#) to block these pages from being indexed.
- **Understand your content management system:** Make sure you're familiar with how content is displayed on your web site. Blogs, forums, and related systems often show the same content in multiple formats. For example, a blog entry may appear on the home page of a blog, in an archive page, and in a page of other entries with the same label.
- **Minimize similar content:** If you have many pages that are similar, consider expanding each page or consolidating the pages into one. For instance, if you have a travel site with separate pages for two cities, but the same information on both pages, you could either merge the pages into one page about both cities or you could expand each page to contain unique content about each city.

Google does not recommend blocking crawler access to duplicate content on your website, whether with a robots.txt file or other methods. If search engines can't crawl pages with duplicate content, they can't automatically detect that these URLs point to the same content and will therefore effectively have to treat them as separate, unique pages. A better solution is to allow search engines to crawl these URLs, but mark them as duplicates by using the `rel="canonical"` link element, the URL parameter handling tool, or 301 redirects. In cases where duplicate content leads to us crawling too much of your website, you can also [adjust the crawl rate setting](#) in Webmaster Tools.

Duplicate content on a site is not grounds for action on that site unless it appears that the intent of the duplicate content is to be deceptive and manipulate search engine results. If your site suffers from duplicate content issues, and you don't follow the advice listed above, we do a good job of choosing a version of the content to show in our search results.

However, if our review indicated that you engaged in deceptive practices and your site has been removed from our search results, review your site carefully. If your site has been removed from our search results, review our [Webmaster Guidelines](#) for more information. Once you've made your changes and are confident that your site no longer violates our guidelines, [submit your site for reconsideration](#).

In rare situations, our algorithm may select a URL from an external site that is hosting your content without your permission. If you believe that another site is duplicating your content in violation of copyright law, you may contact the site's host to request removal. In addition, you can request that Google remove the infringing page from our search results by [filing a request under the Digital Millennium Copyright Act](#).

updated 03/23/2012

Webmaster Tools - Help for other Google products -

©2012 Google

- [Google Home](#) - [Privacy Policy](#) - [Terms of Service](#)